

# Utaskikérdezés alapú célforgalmi mátrix becselő módszer közforgalmú közlekedéshez

Dr. Földes Dávid<sup>1</sup> – Dr. Csonka Bálint<sup>2</sup> – Dr. Csiszár Csaba<sup>3</sup>

Budapesti Műszaki és Gazdaságtudományi Egyetem  
Közlekedésmérnöki és Járműmérnöki Kar  
Közlekedésüzemi és Közlekedésgazdasági Tanszék

<sup>1</sup>telefon: +36 1 463 1977

e-mail: foldes.david@kjk.bme.hu

<sup>2</sup>telefon: +36 1 463 1977

e-mail: csonka.balint@kjk.bme.hu

<sup>3</sup>telefon: +36 1 463 1978

e-mail: csizar.csaba@kjk.bme.hu

**Kivonat:** A közforgalmú közlekedési szolgáltatás optimalizáláshoz elengedhetetlen a mobilitási szokásokat leíró honnan-hová mátrix. A mátrix értékeinek meghatározásához jellemzően keresztmetszeti forgalomszámlálást és utaskikérdezésre alapuló (önbevallásos útinapló, telefonos felmérés, személyes háztartásfelvétel stb.) módszereket alkalmaznak. A mintavétel nagysága jelentősen növeli a költséget és befolyásolhatja az eredmények megbízhatóságát. A kutatási célunk, olyan honnan-hová mátrix számító módszer kidolgozása, aminek bemeneti adatforrása egy online utazásfelvételi kérdőív. A felmérésben résztvevők rögzítik az utazás kezdő és kiinduló pontját, az első és utolsó megállóhelyet, az utazás időpontját, motivációját és gyakoriságát, az átszállások számát és helyét, valamint az életkort és nemüket. A felvétel alapján a közösségi közlekedőkre vonatkozó utazások száma becsülhető körzetekre nemenként és korcsoportonként figyelembe véve az adott körzet lakosságszámát, nem- és korösszetételét. A becsült utazásszám alapján meghatározható a körzetek és megállóhelyek közötti utazásszám. A módszer újszerűsége, hogy a rögzített utazások megvalósulásának valószínűsége figyelembe vehető. Bevezettük a megbízhatósági rátát, amivel alacsony mintaszám esetén a teljes sokaságra és az egyes csoportokra jellemző értékek súlyozhatók. A módszert magyarországi városra alkalmaztuk. A módszer hozzájárul a közforgalmú közlekedés tervezéséhez.

**Kulcsszavak:** honnan-hová mátrix, utazásfelvétel, kérdőív, közforgalmú közlekedés, becselő módszer

## Nomenklatúra

$\alpha$	nem indexe	$n$	mintanagyság
$\beta$	korcsoport indexe	$N$	alapsokaság mérete
$\gamma$	motiváció indexe	$p$	jellemzők megoszlása a teljes mintán
$\sigma$	naptípus indexe	$s$	relatív szórás
$\tau$	időszak	$S$	összes utazás
$\varphi$	napszak indexe	$t$	megbízhatósági szint
$\omega$	utazási mód indexe	$u$	rögzített utazások száma
$a$	becsült utazások százalékos megoszlása	$u'$	fajlagos utazási igény
$c$	célforgalmi mátrix cellája	$U$	utazások összesített száma
$d$	pontosság	$UB$	becsült utazások száma
$E$	célforgalmi egységmatrrix	$UR$	utazások becsült száma
$f$	utazások gyakorisága	$WB$	becsült utazások súlyszáma
$i$	induló körzet indexe	$WR$	rögzített utazások súlyszáma
$j$	érkező körzet indexe		
$k$	utazást rögzítő személy lakó körzet indexe		

## Bevezetés

A helyi közforgalmú közlekedés biztosítása elengedhetetlen nagyobb városokban, hisz a tömeges, egyirányban és időszakban megjelenő utazási igényeket gazdaságosan, hatékonyan és fenntartható módon csak közforgalmú közlekedéssel lehet kiszolgálni [1]. Bár a közforgalmú közlekedés üzemeltetése általában veszteséges, csak a jegyárakból nem lehet fedezni működését [2], a járulékos hasznok jelentősek (alacsonyabb környezetterhelés, torlódások csökkenése, közlekedés kisebb helyfoglalása stb.) [3]. A hatékonyság fokozása és az utaselégedettség növelése érdekében a szolgáltatás folyamatos felülvizsgálata szükséges. A hálózattervezés (viszonylatok, járatszám) az utazási igények ismeretében lehetséges, amit az igény nagysága, térbeni és időbeli megoszlása, motivációja és egyéb jellemzői (pl. utazóicsoporthoz) írják le. Az igények alapján különböző felbontású (idő, motiváció, utazói csoport stb.) körzetekre, illetve megállóhelyekre vonatkozó honnan-hová, vagy másnéven célforgalmi mátrixok képezhetők. Az utazási igényeket indirekt módon a területek vizsgálatával (forgalomvonzó létesítmények, népsűrűség stb.), míg direkt módon utasszámlálással és kikérdezéssel lehetséges. Az utazók kikérdezésével az utazás számos jellemzője meghatározható, ugyanakkor kihívást jelent a reprezentativitás. Továbbá bizonyos igények feltáratlanok maradnak, amelyek becslése szükséges.

A következő kutatási kérdést fogalmaztuk meg: hogyan becsülhetők az utazási igények, hogyan határozhatók meg ezek alapján a célforgalmi mátrixok? Ezek megválaszolásához utaskikérdezés alapú célforgalmi mátrix becslő módszert dolgoztunk ki, amelyben a nem ismert utazások valószínűsége a rögzített utazások megvalósulásának valószínűségével becsülhető.

A cikk felépítése a következő: az 1. fejezetben áttekintjük a releváns szakirodalmak eredményeit. A 2. fejezetben részletezzük a minta nagyság meghatározásához alkalmazott módszert, illetve a kidolgozott célforgalmi mátrix becslő módszert. A 3. fejezetben esettanulmányon keresztül bemutatjuk a módszer alkalmazását. Végezetül a kutatás során levont következtetéseket és a továbbfejlesztés irányait foglaljuk össze.

## 1. Irodalomkutatás

Az utazási igényeket, a célforgalmi mátrixokat direkt módon becslő tanulmányok is két csoportba sorolhatók az alkalmazott megközelítés alapján:

- kikérdezésre, vagy
- forgalomszámlálásra épülő becslések.

Kikérdezésből származó adatok alapján jól becsülhető a célforgalmi mátrix, például Bayesiánus statisztikával [4], [5] és véletlen erdő modellt alkalmazva [6]. A kikérdezéses adatgyűjtés előnye, hogy az utazás egyéb jellemzői is vizsgálhatók, ami a közlekedési döntésmodellezés alapja. Például feltárható az egyes utazói csoportok közlekedési és személyes jellemzői közötti kapcsolat [7]. A népszámlálási adatokat felhasználva vizsgálták a munka motiváltaságú utazásokat, ami alapján új területkategóriák képezhetők [8]. A kikérdezésen alapuló módszerek hátránya az alacsony válaszadási hajlandóság, a hiányos válaszadás, utazói csoport és utazástípusok reprezentativitása, az alul vagy túlbecsült utazási idő ([9],[10]). Ezek miatt megnőtt az igény egyéb adatgyűjtési módszerek iránt.

A forgalomszámlálásra épülő megoldások már az 1970-es években megjelentek [11], de igazán a technológiai változásnak köszönhetően terjedtek el, mivel nagymennyiségű, valós idejű adat lehetővé teszi a dinamikus célforgalmi mátrixok képzését. A kihívást az adatok feldolgozását támogató eljárások fejlesztése jelenti. A gépkocsit használók célforgalmi mátrixa entrópiailag maximalizáló módszerrel meghatározható [12]. Az eljárás hátránya, hogy a bemenő adatot hibamentesnek tekinti. Alkalmazható legkisebb négyzetek módszere és a méhkolónia optimalizáló eljárás is [13], aminek előnye az alacsony számítási idő és a valós idejű forgalommodellezés támogatása. Fuzzy követő algoritmus használatával az egyes szakaszok és irányok közötti relatív kapcsolat megadása nélkül is előállítható a célforgalmi mátrix [14]. A célforgalmi mátrix megbízhatósága fokozható a forgalomszámláló szenzorok optimális elhelyezésével [15]. A forgalomszámlálásra épülő módszerek hátránya, hogy költséges az összes útszakasz figyelembevétele, az utazás és az utazó egyéb jellemzői egyáltalán nem, vagy csak részben mérhetők, illetve a fejlett forgalomszámláló megoldások nem mindenhol elérhetők.

A hagyományos forgalomszámláló megoldásokon túl egyéb adatforrások is megjelentek, például a mobiltelefonok cellainformációi [16] vagy az elektronikus jegyrendszerből származó utazási adatok [17],

[18], [19]. Ugyanakkor az elektronikus jegyrendszerből származó adatok önmagukban nem kellően pontosak, más adatforrások figyelembevétele is indokolt [20].

Amennyiben a rendelkezésre álló pénzügyi forrás alacsony, a kikérdezéses felmérések a hatékonyak, ellenkező esetben a forgalomszámlálásra épülő megoldások [14]. Azonban a kikérdezést és forgalomszámlálást ötvöző megoldások is megtalálhatók az irodalomban. Például a forgalomszámlálásból származó adatok alapján a kikérdezéssel meghatározott utazásszámok frissíthetők, pontosíthatók [21], [22], [23], [24]. Továbbá a közösségi közlekedésre érvényes célforgalmi mátrix meghatározásához kidolgozott iteratív becslő módszer jóságát nem befolyásolja a közlekedési hálózat nagysága [24]. A modell újszerűsége, hogy a szakaszok jellemzőinek a figyelembevételével az utazói döntések is vizsgálhatók. Ötvözhető a jegyértékesítésből, forgalomszámlálásból és utas kikérdezésből származó adatok is; például a vasúti közlekedés célforgalmi mátrixának meghatározásához [25].

Az irodalomkutatás alapján megállapítottuk, hogy a részletes vizsgálatokhoz elengedhetetlenek a kikérdezéses adatgyűjtést, azonban a célforgalmi mátrixok megbízhatóságát jelentősen javíthatja a forgalomszámlálási adatok figyelembevétele.

## 2. Módszertan

A kidolgozott célforgalmi mátrix becslő módszer részletes utaskikérdezésen alapul.

### 2.1 Mintanagyság meghatározása

Mintavételes forgalomfelvételnél megkülönböztethető:

- véletlenszerű mintavétel,
- szisztematikus mintavétel és
- rétegzett mintavétel.

Feltételezzük, hogy ismert a teljes alapsokaság mérete, illetve az alapsokaságon belül bizonyos jellemzők szerint (pl.: nem, korcsoport) képzett homogénnek tekinthető rétegek nagysága, ezért a rétegzett mintavétel alkalmazható, ami a valóságot jobban leképező minta kialakításához vezet. Adottnak tekintjük a lakosság nem és korcsoport szerinti megoszlását egy  $k$  lakóközretnen belül. A szükséges mintanagyságokat nemre és korcsoportra, mint homogén rétegekre határozzuk meg.

A kérdőíves kikérdezés során az általános személyes jellemzők (nem, korcsoport) mellett tetszőleges számú jellegzetes utazás felvehető. Rögzítendő a gyakoriság, motiváció, jellemző naptípus és napszak, közlekedési mód, indulási és érkezési helyszín. A motivációt csak az odautazásoknál rögzítettük. Közforgalmú közlekedés esetén rögzíthető az induló és érkező megállóhely. Számos további jellemző is felmérhető (használt viszonylat, jegytípus, átszállás helye stb.) további elemzések készítéséhez, amelyek a célforgalmi mátrixok elkészítését nem befolyásolják.

Használt indexek:

- $k$  lakóközretn  $k \in \mathbb{N}$ ,
- $\alpha$  nem  $\alpha=1..2$  (1: férfi, 2: nő),
- $\beta$  korcsoport  $\beta=1..8$  (1: 0-6, 2: 7-14, 3: 15-18, 4: 19-24, 5: 25-39, 6: 40-54; 7: 55-64, 8: 65+)
- $\gamma$  motiváció  $\gamma=1..9$  (1: munkába utazás, 2: iskolába utazás, 3: munkavégzés, 4: vásárlás, 5: ügyintézés, 6: egészségügy, 7: rekreáció, 8: egyéb).

*Feltételezzük, hogy statisztikai, népszámlálási adatok alapján a lakóközretnek lakosságszáma nem és korcsoport szerint elérhető.*

Véges alapsokaság esetén a szükséges elméleti mintanagyságot  $k$  közretnenként az (1) alapján az általános statisztikából ismert egyenlet alapján határozható meg.

$${}^k n = \frac{{}^k N \cdot t^2 \cdot s^2}{t^2 \cdot s^2 + (N_i - 1) \cdot d^2} \quad (1)$$

Ahol:

- $N$  alapsokaság mérete (lakosságszám),
- $t$  megbízhatósági szint
- $s$  relatív szórás,
- $d$  pontosság.

Normális eloszlást feltételezve a 95%-os megbízhatósági szinthez a következő megbízhatósági szint értékek tartoznak az alapsokaság méretének függvényében (standard normális eloszlás táblázatból):

- $500 < N$  esetén  $t = 1,96$ ,
- $200 < N \leq 500$  esetén  $t = 1,9712$ ,
- $100 < N \leq 200$  esetén  $t = 1,9840$ .

A relatív szórás a jelenség, vagyis az utazási szokások stabilitását írja le. Minél alacsonyabb a szórás, annál nagyobb megbízhatósággal következik be a rögzített utazás, vagyis stabil. Értéke 0,2-0,5 között változik. Az utazási motivációkhoz az 1. táblázatban leírt relatív szórás értékeket társítottuk. A visszautazáshoz tartozó relatív szórás érték megegyezik az odaútas motivációjához tartozó relatív szórás értékkel.

1. táblázat Relatív szórás értékei motivációnként

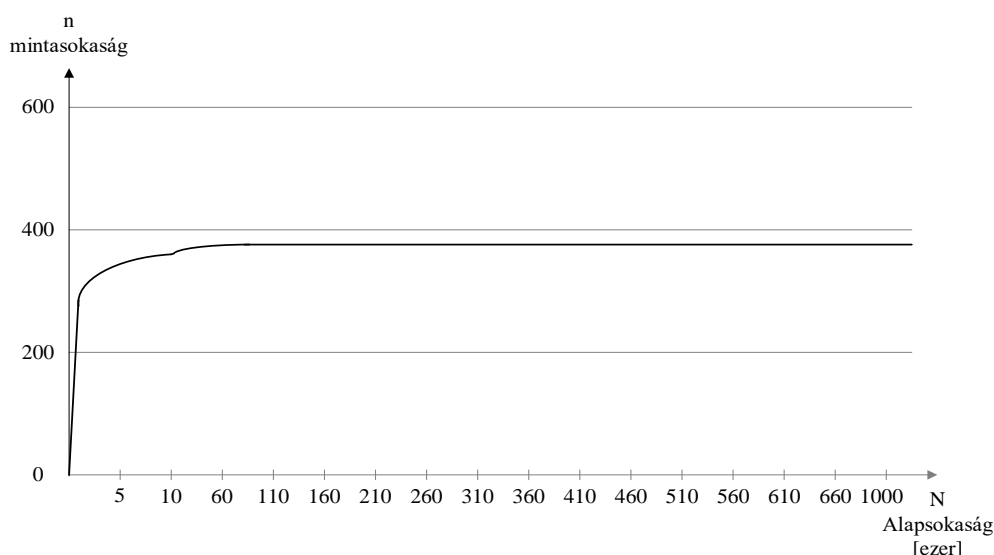
$\gamma$	Motiváció	Relatív szórás $s''$
1-3	hivatásforgalom (munkába, iskolába, munkavégzés)	0,2
4-5	vásárlás és ügyintézés	0,3
6	egészségügy	0,4
7-8	rekreáció és egyéb	0,5

Kifejezi a mintavételi hibáját, az abból eredő hibát, hogy nem az egész sokaság került megfigyelésre. A becslés és az ismeretlen sokaság hibaérték közötti különbséget jelenti. Értéke az alapsokaság méretétől függ (2. táblázat).

2. táblázat Pontosság értékek

Alapsokaság $N$ [fő]	Pontosság $d$
0-20	0,4
20-50	0,3
50-100	0,2
100-500	0,1
>500	0,05

A minta nagyságát az alapsokaság nagyságának a függvényében mutatja be az 1. ábra 95%-os megbízhatósági szint és 5%-os pontosság mellett [26]. Pár ezernél nagyobb alapsokaság felett a szükséges mintanagyság érdemben nem változik. Az értéket a stabilitás tényező értékének megválasztása befolyásolja.



1. ábra: Alapsokaság és a mintasokaság összefüggése (forrás [26])

A  $k$  körzetben szükséges minta nagyságát különböző jellemzőnként (rétegenként) és a rétegek kombinációjával is meg lehet adni, például  $\alpha$  nemenként és  $\beta$  korcsoportonként (2). Amennyiben a jellemzők közül valamelyik szerint nem történik csoportosítás, úgy az adott index értéke 0.

$${}^k n^{\alpha, \beta} = {}^k p^{\alpha, \beta} \cdot n_i \quad (2)$$

A számított értékek különböző  $x$  jellemzőkre, szűrések elvégzésével pontosíthatók  $x = \{\alpha, \beta, \lambda, \sigma, \varphi, \omega\}$ .

### 1. lépés: Rögzített utazások számának meghatározása

A  $k$  lakókörzetenként az összes  $i$  körzetből  $j$  körzetbe végzett **utazások számának meghatározása**  $\tau$  időszakra, valamint  $x$  jellemzőre és azok kombinációjára vonatkozóan:  ${}^k u(\tau)_{i,j}^x$ . Egy rögzített utazáshoz tartozó utazásszám az  $f$  gyakoriság érték reciproka. Az alkalmazott utazási gyakoriság kategóriákat és a hozzá tartozó  $f$  értéket az 3. táblázat tartalmazza,  $\tau=1$  nap.

3. táblázat Utazási gyakoriság értékek

Gyakoriság kategóriák	$u(\tau=1)$
naponta	1
hetente 2-3 alkalommal	1/3
hetente	1/7
két hetente	1/14
ritkábban	1/28

### 2. lépés: Fajlagos utazási igény számítása [igény/nap]

A fajlagos utazási igény kifejezi egy átlagos személy átlagos napi utazásainak számát. Tehát ez a mutató nem körzetekre, hanem személyes jellemzőkre határozható meg,  $x=\{\alpha, \beta\}$ . Számítása: összes rögzített utazások számának és az összes megkérdezett számának hányadosa (3).

$$u'(\tau)^x = \frac{\sum_{i,j} {}^k u(\tau)_{i,j}^x}{n^x} \quad (3)$$

### 3. lépés: Célforgalmi egységvektor számítása [%]

A  $j$  célkörzetbe utazás valószínűségének meghatározása  $x$  jellemzőre (4).

$$E(\tau)_j^x = \frac{{}^k u(\tau)_{i,j}^x}{\sum_j {}^k u(\tau)_{i,j}^x} \cdot 100 \quad (4)$$

### 4. lépés: Körzetben élők összes utazása

A  $k$  körzetben élők összes utazásának meghatározása  $\tau$  időszakra  $\alpha$  nemre és  $\beta$  korcsoportra vonatkozó lakosságszám és fajlagos utazási igény szorzatának összegzésével (5). (A körzetben élők száma nem és korcsoport szerinti bontásban általában elérhető, így a körzetben élők összes utazását ezen jellemzők alapján részletezhető.)

$${}^k S(\tau)^x = {}^k N^x \cdot u'(\tau)^x \quad x=\{\alpha, \beta\} \quad (5)$$

### 5. lépés: Utazások becsült száma

Az utazások becsült számának meghatározása  $\tau$  időszakra és  $k$  lakókörzetre vonatkozóan (6).

$${}^k UR(\tau)^{\alpha, \beta, \gamma, \sigma, \varphi, \omega} = {}^k S(\tau)^{\alpha, \beta} \cdot \frac{{}^k u(\tau)_{i,j}^{\alpha, \beta, \gamma, \sigma, \varphi, \omega}}{{}^k u(\tau)_{i,j}^{\alpha, \beta}} \quad (6)$$

A rögzített utazások számára igaz, hogy  $u(\tau)_{i,j}^{\alpha, \beta, \gamma, \sigma, \varphi, \omega} \leq u(\tau)_{i,j}^{\alpha, \beta}$ . Ha csak  $\alpha$  nemre és  $\beta$  korcsoportra vonatkozó szűrő feltételek kerültek beállításra, akkor  ${}^k UR(\tau)^{\alpha, \beta, \gamma, \sigma, \varphi, \omega} = {}^k S(\tau)^{\alpha, \beta}$ .

Az  $x$  jellemzők, vagyis  $\alpha$  nem,  $\beta$  korcsoport,  $\gamma$  motiváció,  $\sigma$  naptípus,  $\varphi$  napszak,  $\omega$  utazási mód jellemzők külön jelölésétől a továbbiakban eltekintünk.

### 6. lépés: Megbízhatósági ráta meghatározása

A megbízhatósági rátát  $k$  lakókörzetenként határoztuk meg. A megbízhatósági ráta megadásához a megbízhatósági szint számítása szükséges, a mintanagyság meghatározásához bevezetett (1) egyenlet átrendezésével (7).

$${}^k t = \sqrt{\frac{{}^k n \cdot ({}^k N - 1) \cdot d^2}{({}^k N - {}^k n) \cdot s^2}} \quad (7)$$

A  $d$  pontosság a 2. táblázat alapján határozható meg a teljes mintanagyságot figyelembe véve. A teljes mintára vonatkozó relatív szórás ( $s$ ) a megvalósult utazások  $\gamma$  motivációnként megoszlása és a motivációkra jellemző relatív szórás (1. táblázat) alapján számítható (8).

$$s = \frac{u(\tau)_{i,j}^\gamma}{\sum_{\gamma} u(\tau)_{i,j}^\gamma} \cdot s^\gamma \quad (8)$$

Adott megbízhatósági szint tartományokhoz diszkrét értéket rendeltünk (4. táblázat).

4. táblázat Megbízhatósági ráta értéke

$t$	Megbízhatósági szint
0-1,2816	90% vagy nagyobb
1,2816-1,6449	80-90% közötti
>1,6449	80% alatti

#### 7. lépés: Rögzített utazások súlysza

A megbízhatósági szint alapján meghatározható, hogy egy  $k$  körzet utazásai milyen arányban becsülhetők a  $k$  körzetben élők által rögzített utazások és a teljes vizsgált területen élők által rögzített utazások alapján. Az adott lakókörzetben élők által rögzített utazások alapján becsült utazások aránya a következő:

- $a(t)=0,9$ , vagyis 90%, ha a megbízhatósági ráta 90% vagy magasabb,
- $a(t)=0,8$ , vagyis 80%, ha a megbízhatósági ráta 80-90%,
- $a(t)=0,5$ , vagyis 50%, ha a megbízhatósági ráta 80% alatti.

A  $k$  lakókörzetenként a rögzített utazásokhoz tartozó súlysza számot a  $k$  lakókörzetben élők összes utazása és a rögzített utazások száma alapján határoztuk meg (9). A körzeten belüli összes utazás bizonyos részét a felvett utazások alapján becsüljük a későbbi lépésekben.

$${}^kWR(\tau) = \frac{{}^kUR(\tau) \cdot {}^ka(t)}{\sum_i {}^ku(\tau)_{i,j}} \quad (9)$$

Azon körzetekben, ahol kevés utazás került rögzítésre, azért, a körzet felülreprezentálását elkerülve, a rögzített utazások súlysza nem lehet nagyobb, mint 50% ( $a(t)=0,5$ ). Ebben az esetben a becsült utazások részaránya több, mint 50%.

Ezek alapján számítható az utazások korrigált becsült száma, vagyis a rögzített utazások alapján számított utazások nagysága az összes utazáson belül (10).

$${}^kUR(\tau)' = {}^ku(\tau)_{i,j} \cdot {}^kWR(\tau) \quad (10)$$

Tekintsük a következő példát:  $k$  lakókörzeten belüli összes utazás száma  ${}^kUR(\tau)=1000$ , a megbízhatósági ráta 80% ( ${}^ka(t)=0,8$ ), a körzetben rögzített, vagy a kitöltők által felvett utazások száma  ${}^ku(\tau)_{i,j}=125$ . Ez esetben az utazások 80%-át a rögzített utazások alapján becsültük, azaz  ${}^kUR'=800$  utazást. A rögzített utazásokat megszoroztuk  $800/125=6,4$ -el. Tehát  ${}^kWR(\tau)=6,4$ .

#### 8. lépés: Becsült utazások számának meghatározása

A maradék utazás az összes  $x$  szűrő feltételnek megfelelő többi utazás alapján becsülendő. Ez alapján meghatározható minden  $k$  körzetben élők azon utazásainak száma, amelyek az összes körzet lakói által rögzített utazásai alapján becsülendők. A becsült utazások száma a  $k$  körzet lakóinak összes utazása, valamint a rögzített utazások számának és a hozzátartozó súlysza szorzatának különbsége (11).

$${}^kUB(\tau) = {}^kS(\tau) - {}^kUR(\tau) \cdot {}^kWR(\tau) \quad (11)$$

#### 9. lépés: Célforgalmi mátrixok számítása

##### Körzet-körzet célforgalmi mátrix

A  $k$  lakókörzetben élő utazók körzet-körzet közötti utazások összesített száma: a rögzített utazások és a rögzített utazások súlysza szorzata, valamint  $k$  körzetenként a becsült utazások és a célforgalmi egységvektor szorzata. Külön kezelendő az adott  $k$  körzeten belül végrehajtott utazások (12), illetve az eltérő kiinduló és érkező körzettel rendelkező utazások (13). A 2-vel osztás amiatt szükséges, mert a

becsült utazások fele a célforgalmi egységvektor szerinti körzetekbe indul, a másik fele azokból a körzetekből érkezik. A rögzített utazásokat egységesen megszoroztuk ugyanazzal a  $WR(\tau)_i$  értékkel egy körzeten belül, hogy a megszorozott utazások száma megegyezzen a megbízhatósági ráta alapján számított százalékos értékkel súlyozott összes utazásszámmal.

$${}^kU(\tau)_{k,k} = {}^k u(\tau)_{k,k} \cdot {}^kWR(\tau) + {}^kUB(\tau) \cdot E(\tau)_k \quad (12)$$

$${}^kU(\tau)_{i,j} = {}^k u(\tau)_{i,j} \cdot {}^kWR(\tau) + \frac{{}^kUB(\tau)}{2} \cdot E(\tau)_j \quad i, j \neq k \quad (13)$$

Példa számítás: tekintsük a következő három körzetet: A, B, C. A  $k=A$  körzet lakosainak száma  ${}^AN=10$  fő. A megbízhatósági ráta 90% ( $a(t)=0,9$ ). A fajlagos utazási igény mindhárom körzetet figyelembevéve  $U(\tau)=2$ . Vagyis a  $k=A$  körzetben lakók becsült összes utazásszáma  ${}^AS(\tau)=20$ . Tételezzük fel, hogy a következő 6 utazás került rögzítésre  ${}^k u(\tau)_{i,j}$ :

- $A \rightarrow A$ :  $u(\tau)_{A,A}=0$ ,
- $A \rightarrow B$ :  $u(\tau)_{A,B}=2$ ,  $u(\tau)_{B,A}=2$  (oda-vissza),
- $A \rightarrow C$ :  $u(\tau)_{A,C}=1$ ,  $u(\tau)_{C,A}=1$  (oda-vissza).

A  $k=A$  körzetre vonatkozó számítás:

- Mivel a megbízhatósági ráta 90%, ezért az utazások 90%-át,  ${}^AUR(\tau) = 18$ -at, a rögzített utazások alapján becsülünk.
- A rögzített utazások súlyszáma  ${}^AWR(\tau) = 18/6=3$ .
- A becsült utazások száma  ${}^AUB(\tau)=2$ .
- Tételezzük fel, hogy a célforgalmi egységvektor  $E(\tau)_j$  alapján az utazások 25%-25%-a az A és a C körzetbe érkezik, 50%-a a B körzetbe.

Ezek alapján az utazások száma:

- $U(\tau)_{A,A} = 0,25 \cdot 2 = 0,5$ , mivel nem volt rögzített utazás, így csak a célforgalmi egységvektor alapján becsült utazás szerepel,
- $U(\tau)_{A,B} = 3 \cdot 2 + 2/2 \cdot 0,5 = 6,5$ ,
- $U(\tau)_{B,A} = 3 \cdot 2 + 2/2 \cdot 0,5 = 6,5$ , az A-B irány szimmetrikus, mivel a rögzített utazások száma is megegyezik mindkét irányban.
- $U(\tau)_{A,C} = 3 \cdot 1 + 2/2 \cdot 0,25 = 3,25$ ,
- $U(\tau)_{C,A} = 3 \cdot 1 + 2/2 \cdot 0,25 = 3,25$ , az A-C irány is szimmetrikus.

Így a  $k=A$  körzetben lakók utazásainak összesített száma:  ${}^kU(\tau) = {}^kU(\tau)_{k,k} + \sum_k {}^kU(\tau)_{i,j} = 20$

Ezek alapján a célforgalmi mátrix celláinak értéke az egyes körzetekben élők utazásainak az összesítésével számítható (13).

$$c_{i,j} = \sum_k {}^kU(\tau)_{i,j} \quad (14)$$

#### Megállóhely-megállóhely célforgalmi mátrix

A módszer alkalmas megállóhely-megállóhely közötti célforgalmi mátrixok előállítására is. A lépések megegyeznek, ugyanakkor új indexek bevezetése szükséges. Az  $i$  és  $j$  körzet-indexek helyett  $v$  és  $w$  megállóhely-indexek használata szükséges. Például az utazások száma  $v$  és  $w$  megállóhely között  $\tau$  időszakra:  ${}_{v,w}U(\tau)$ .

#### Körzet-megállóhely célforgalmi mátrix

A módszer alkalmas a körzet-megállóhely vegyes célforgalmi mátrixok előállítására. Ez a megállóhely-megállóhely célforgalmi mátrixon alapul; az egyes megállóhelyek körzetekbe sorolása szükséges. Jelöléstechnika példa: az utazások száma  $i$  körzet és  $q$  megállóhely között  $\tau$  időszakra:  ${}_{v,w}U(\tau)_{i,j}$ , ahol  $v=0$ ,  $w=N^+$ ,  $i=N^+$ ,  $j=0$ , vagyis az induló megállóhely és érkező körzet 0 értéket kap.



### 3. Esettanulmány

A módszert egy magyarországi megyei jogú városra alkalmaztuk. A példaterület kiterjedt közforgalmú közlekedési hálózattal, autóbusz-, trolibusz- és villamoshálózattal rendelkezik.

Célunk a közforgalmú közlekedés honnan-hová mátrixának elkészítése volt, így az alapsokaságként nem a teljes lakosságot, hanem a városban a közforgalmú közlekedésre jellemző modal-share értékkel csökkentett lakosság számot vettük alapul (a közforgalmú közlekedés modal-share értéke 40%). Feltételeztük, hogy a nem, korcsoport és lakóhely nem befolyásolja jelentősen a közösségi közlekedés részarányát, ezért az alapsokaság a szűrés utáni lakosság szám 40%-a.

#### 3.1 Mintanagyság

A városi körzetekre elérhető nemre és korcsoportra is lebontott lakosság szám adatok alapján (adatok forrása: önkormányzat) meghatároztuk a körzetekben szükséges minta nagyságát  $s=0,2$  (stabil igények) és  $s=0,5$  (instabil igények) stabilitási értékek mellett. A mintavétel nagyságát nemekre vonatkozóan határoztuk meg. Az így kapott mintanagyságot lakosság arányosan osztottuk szét a korcsoportok között. Az 5. táblázat az  $N$  lakosság számot (alapsokaság) nemenként, férfiak esetében korcsoporti bontásban, és a szükséges  $n$  minta nagyságát mutatja be két stabilitási szint esetén három mintakörzetre:

- 1 nagy lakótelep,
- 2 közepes lakótelep,
- 3 rekreációs terület.

5. táblázat: Szükséges mintanagyság - kivonat

k	$\alpha$	férfi									nő	$\Sigma$
		1									2	
	$\beta$	0-6	7-14	15-18	19-24	25-39	40-54	55-64	65+	$\Sigma$	$\Sigma$	
		1	2	3	4	5	6	7	8			
1	N	285	206	321	1064	1157	506	713	353	4605	5556	10160
	n (s=0,2)	2	1	2	6	7	3	4	2	28	33	61
	n (s=0,5)	11	8	12	40	43	19	27	13	172	206	378
2	N	59	45	64	212	246	104	142	78	950	1120	2070
	n (s=0,2)	2	1	2	6	7	4	4	3	29	32	61
	n (s=0,5)	10	8	11	37	42	18	25	13	164	194	358
3	N	6	6	2	17	22	13	15	6	86	97	183
	n (s=0,2)	1	0	0	1	2	1	1	0	7	8	15
	n (s=0,5)	3	2	1	7	9	6	6	3	37	43	80

A szükséges minta nagysága a legtöbb körzetben hasonló mértékű. Ennek oka, hogy 500 fős alapsokaság (lakosság szám) felett a 95%-os megbízhatósághoz tartozó szükséges mintanagyság közel állandónak tekinthető; a mértéke a stabilitás értékétől függ csak.

#### 3.2 Célforgalmi mátrix becslése

A becslési módszer alapjául szolgáló kérdőíves kutatás 2020-ban készült. Összesen 5452 személy által végrehajtott 9648 egyedi utazás került rögzítésre.

##### 1. lépés: Rögzített utazások száma

A példakörzetek tényleges mintanagyságát a 6. táblázat tartalmazza, férfiak esetében korcsoport szerinti bontásban is. A körzetek többségében a női válaszadók voltak többségében; ez összhangban van azzal, hogy az alapsokaságban is a nők aránya a nagyobb.

6. táblázat: Tényleges mintanagyság - kivonat

körzet	férfi									nő	össz.
	0-6	7-14	15-18	19-24	25-39	40-54	55-64	65+	össz.	össz.	
1	0	14	23	52	102	66	17	22	296	432	728
2	0	4	2	3	14	18	0	0	41	53	94
3	0	0	0	0	1	1	0	0	2	4	6

A válaszadók rögzített utazásai  $\tau=1$  napra vonatkozóan a mintakörzetekben:  $k=1$ : 1147,  $k=2$ : 161,  $k=3$ : 12 [utazás/nap]

## 2. lépés: Fajlagos utazási igény

A napi fajlagos utazás számát nem és korcsoport szerinti bontását a 7. táblázat tartalmazza. A fajlagos utazási igény értéke 1,7 azokban a csoportokban, ahol a megkérdezettek száma kevesebb, mint 50 fő. Összesen 4 csoportban volt kevesebb 50 főnél a megkérdezettek száma (férfi/nő 0-6 éves korig, 7-14 éves korig). A fajlagos utazási igény a 15-18 és a 40-54 korosztályban a legmagasabb, míg a 65+ korosztályban a legalacsonyabb.

7. táblázat: Fajlagos utazási igény nemenként és korcsoportonként

			$\beta$							
			1	2	3	4	5	6	7	8
			0-6	7-14	15-18	19-24	25-39	40-54	55-64	65+
$\alpha$	1	férfi	1,7	1,7	1,86	1,70	1,78	1,73	1,64	0,99
	2	nő	1,7	1,7	1,72	1,75	1,70	1,82	1,76	0,90

## 3. lépés: Célforgalmi egységvektor

A célforgalmi egységvektort, vagyis az utazások egy adott körzetbe végződésének arányát (kitöltők utazásai alapján) néhány példa körzetre a 8. táblázat mutatja be. A táblázatban a körzetek rövid jellemzése is olvasható.

8. táblázat: Célforgalmi egységvektor – kivonat

$j$	Megnevezés	$E_j$ [%]
1	nagy lakótelep	12,27
2	közepes lakótelep	1,48
3	rekreációs terület a város szélén 1	0,19
4	belváros	18,66
5	kertváros 1	3,90
6	iparterület 1	8,83
7	egyetemi terület	4,75
8	rekreációs terület a város szélén 2	1,95
9	iparterület 2	4,71
10	kertváros 2	2,36

## 4. lépés: Körzetben élők összes utazása

A  $k$  mintakörzetek összes napi utazásszáma [utazás/nap]:  $k=1$ : 17033,  $k=2$ : 3459,  $k=3$ : 305 [utazás/nap].

## 5. lépés: Utazások becsült száma

A szűrési feltételeket csak  $\alpha$  nemre és  $\beta$  korcsoportra vonatkozóan végeztük el, így az utazások becsült száma megegyezik a körzetben élők összes utazásával. Bár a naptípusra, napszakra, motivációra vonatkozó paraméterek is csökkentik a mintavétel nagyságát, de az alapsokaságról nincs elérhető adat. Így például nem mondható meg a 65 év feletti férfiak 6 és 9 óra közötti munka motivált utazásainak a megbízhatósága.

## 6. lépés: Megbízhatósági ráta

A megbízhatósági szint számításához meghatároztuk az utasok utazásainak  $\gamma$  motiváció szerinti megoszlását a teljes mintára:

$\gamma=1-3$  hivatásforgalom (munkába, iskolába, munkavégzés): 67%,

$\gamma=4, 5$  vásárlás és ügyintézés: 20%,

$\gamma=6$  egészségügy: 3%,

$\gamma=7, 8$  rekreáció és egyéb: 10%.

Ezek alapján a mintavételre érvényes utazásszámmal súlyozott átlagos szórás értéke  $s=0,2565$ .

A mintavétel megbízhatósági szintjét körzetek szerinti (9. táblázat) és a körzeteken belül korcsoportok szerinti (10. táblázat) bontással is meghatároztuk. Az összes körzet 73%-ban a megbízhatósági szint 90% vagy magasabb, csupán 11%-ban 80% alatti. Ezen körzetek gyéren lakott, alacsony beépítettségű, peremterületek. Korcsoport-szerinti bontás alapján a 0-6 és az 55-64 korosztály volt alul reprezentálva a teljes

városra vonatkozóan, ezen korcsoportoknál a megbízhatósági szint 80% alatti. A 15-19, a 25-39, illetve különösen a 40-54 korcsoportra vonatkozó megbízhatósági szint 90%-nál magasabb a körzetek többségében. Bár a  $k=2$  körzetben csupán a 40-54 korosztály esetében magas a megbízhatósági szint, ezen korosztály magas alapsokaság értéke miatt a teljes körzet megbízhatósági szintje is 90%-nál magasabb.

9. táblázat: A mintavétel megbízhatósági szintje az egyes körzetekre

$k$	$t$	Megbízhatósági szint
1	5,8987	90% vagy magasabb
2	2,2959	90% vagy magasabb
3	1,1643	80% alatti

10. táblázat: A mintavétel megbízhatósági szintje korcsoport és lakókörzeti felbontásban

		$\alpha$							
		1	2	3	4	5	6	7	8
		0-6	7-14	15-18	19-24	25-39	40-54	55-64	65+
$k$	1	<80%	>=90%	<80%	>=90%	>=90%	>=90%	>=90%	<80%
	2	<80%	80-90%	<80%	80-90%	<80%	>=90%	<80%	<80%
	3	<80%	<80%	<80%	<80%	<80%	>=90%	<80%	<80%

#### 7. lépés: Rögzített utazások súlyszáma

Mintakörzetenként a rögzített utazások súlyszáma, a következők:  $k=1$ : 13,37,  $k=2$ : 19,31,  $k=3$ : 12,66.

Tehát a  $k=1$  körzet esetében a rögzített utazásokat a súlyszámmal megszorozva számítható a rögzített utazások alapján becsült utazások száma:  ${}^1UR = {}^1u \cdot {}^1WR = 1147 \cdot 13,37 = 15335$  utazás/nap.

#### 8. lépés: Becsült utazások száma

Mintakörzetenként az összes körzet lakói által rögzített utazásai alapján becsülendő utazások száma:  $k=1$ : 1703,  $k=2$ : 346,  $k=3$ : 153 [utazás/nap].

#### 9. lépés: Célforgalmi mátrix

A célforgalmi mátrixot  $k$  minta- és néhány további példakörzetre a 11. táblázatban foglaltuk össze.

11. táblázat Körzet-körzet célforgalmi mátrix - kivonat

$i \backslash j$	1	2	3	4	5	6	7	8	9	10	...	összesen
1	1164	157	30	2311	530	596	112	310	541	455	...	11308
2	61	7	14	508	89	197	9	6	57	7	...	1907
3	12	14	1	54	3	10	4	2	4	5	...	239
...	...	...	...	...	...	...	...	...	...	...	...	...
összesen	11026	1811	223	15665	4123	8333	4071	1842	4621	2675	...	102227

Minden körzetből erős a forgalom a belvárosi területre ( $j=4$ ), valamint az iparterületre ( $j=6$ ). A nagy lakótelepi városrészen ( $i=1$ ) a körzeten belüli utazások száma is magas szemben a kisebb laksűrűségű körzetekben (pl.  $i=2$ ).

## Konklúzió

A kutatás fő eredménye egy utasfelvételen alapuló célforgalmi mátrix becslő módszer. A módszer újszerűsége, hogy a mintavétel megbízhatósága alapján súlyoztuk egy adott körzetben lakók utazásait és az összes megkérdezett által rögzített utazásokat. Minél nagyobb a mintavétel megbízhatósága, az összes utazás annál nagyobb része kerül meghatározásra az adott körzetben rögzített utazások alapján. Ezzel a módszerrel az alul reprezentált körzetekre is nagyobb megbízhatóságú eredmények kaphatók, így a célforgalmi mátrix jósága javul. A módszert egy magyarországi megyei jogú városra alkalmaztuk. A módszer során kihívást jelentett a becslési arányok meghatározása. A módszer limitációja a kérdőíves felmérés; a válaszadók nem feltétlenül rögzítik az összes releváns utazásukat, valamint többnyire a jelenlegi hálózaton végrehajtott utazásukat veszik csak figyelembe. Így a látens, a jelenlegi hálózaton nem megvalósítható utazásokat nem adják meg. A kutatást a módszer validálásával folytatjuk, azaz forgalmi adatok felhasználásával vizsgáljuk a módszerrel becsült utazási számok jóságát.

## Irodalomjegyzék

- [1] Lee S. – Bencekri M.: Urban form and public transport design, Urban Form and Accessibility, Elsevier, 2021 289-306 ISBN 9780128198223 <https://doi.org/10.1016/B978-0-12-819822-3.00018-3>.
- [2] Stone J. – Murphy J.C.: Financing the expansion of mass transit services, Handbook of Sustainable Transport, Elgaronline, 2020 378-387 ISBN: 9781789900460 <https://doi.org/10.4337/9781789900477>
- [3] Stanley J. – Stanley J.: Workshop 7 report: Assessing the wider benefits of public transport projects, Research in Transportation Economics 2020/83 100914 <https://doi.org/10.1016/j.retrec.2020.100914>
- [4] Perrakis K. – Karlis D. – Cools M. – Janssens D. – Vanhoof K. – Wets G.: A Bayesian approach for modeling origin–destination matrices, Transportation Research Part A: Policy and Practice 2012/46 200-212 <https://doi.org/10.1016/j.tra.2011.06.005>
- [5] Perrakis K. – Karlis D. – Cools M. – Janssens D.: Bayesian inference for transportation origin–destination matrices: the Poisson–inverse Gaussian and other Poisson mixtures, Journal of the Royal Statistical Society: Series A (Statistics in Society) 2014/78 271-296.
- [6] Saadi I. – Mustafa A. – Teller J. – Cools M.: A bi-level Random Forest based approach for estimating O-D matrices: Preliminary results from the Belgium National Household Travel Survey, Transportation Research Procedia, 2017/25 2566-2573 <https://doi.org/10.1016/j.trpro.2017.05.301>
- [7] Su R. – McBride E.C. – Goulias K.G.: Unveiling daily activity pattern differences between telecommuters and commuters using human mobility motifs and sequence analysis, Transportation Research Part A: Policy and Practice 2021/147 106-132 <https://doi.org/10.1016/j.tra.2021.03.002>
- [8] Martin D. – Gale C. – Cockings S. – Harfott A.: Origin-destination geodemographics for analysis of travel to work flows, Computers, Environment and Urban Systems 2018/67 68-79 <https://doi.org/10.1016/j.compenvurbsys.2017.09.002>
- [9] Stopher P.R.- Greaves S.P.: Household travel surveys: Where are we going? Transportation Research Part A: Policy and Practice 2007/41 367-381 <https://doi.org/10.1016/j.tra.2006.09.005>
- [10] Sammer G. – Gruber C. – Roeschel G. – Tomschy R. – Herry M.: The dilemma of systematic underreporting of travel behavior when conducting travel diary surveys – A meta-analysis and methodological considerations to solve the problem, Transportation Research Procedia 2018/32 649-658 <https://doi.org/10.1016/j.trpro.2018.10.006>
- [11] Robillard P.: Estimating the O-D matrix from observed link volumes, Transportation Research 1975/9 123-128 [https://doi.org/10.1016/0041-1647\(75\)90049-0](https://doi.org/10.1016/0041-1647(75)90049-0)
- [12] Gonzalez-Calderon C. A. – Posada-Henao J. J. – Restrepo-Morantes S.: Temporal origin–destination matrix estimation of passenger car trips. Case study: Medellin, Colombia, Case Studies on Transport Policy 2020/8 1109-1115 <https://doi.org/10.1016/j.cstp.2020.07.010>
- [13] Caggiani L. – Dell’Orco M. – Marinelli M. – Ottomanelli M.: A Metaheuristic Dynamic Traffic Assignment Model for O-D Matrix Estimation using Aggregate Data, Procedia - Social and Behavioral Sciences 2012/54 685-695 <https://doi.org/10.1016/j.sbspro.2012.09.786>
- [14] Simonelli F. – Marzano V. – Papola A. – Vitiello I.: A network sensor location procedure accounting for o–d matrix estimate variability, Transportation Research Part B: Methodological 2012/46 1624-1638 <https://doi.org/10.1016/j.trb.2012.08.007>
- [15] Reddy K.H. – Chakroborty P.: A fuzzy inference based assignment algorithm to estimate O-D matrix from link volume counts, Computers, Environment and Urban Systems 1998/22 409-423 [https://doi.org/10.1016/S0198-9715\(98\)00037-4](https://doi.org/10.1016/S0198-9715(98)00037-4)
- [16] Bonnel P. – Fekih M. – Smoreda Z.: Origin-Destination estimation using mobile network probe data, Transport Research Procedia, 2018/32 69-81 <https://doi.org/10.1016/j.trpro.2018.10.013>
- [17] Rahbar M. – Hickman M. – Mesbah M. – Tavassoli A.: Calibrating a Bayesian Transit Assignment Model Using Smart Card Data, IEEE Transactions on Intelligent Transportation Systems 2019/20 1574-1583 <https://doi.org/10.1109/TITS.2018.2852726>
- [18] Morency C. – Trépanier M. – Agard B.: Measuring transit use variability with smart-card data, Transport Policy 2007/14 193-203 <https://doi.org/10.1016/j.tranpol.2007.01.001>
- [19] Alsger A. – Mesbah M. – Ferreira L. – Safi H.: Use of Smart Card Fare Data to Estimate Public Transport Origin–Destination Matrix, Transportation Research Record 2015/2535 88-96 <https://doi.org/10.3141/2535-10>
- [20] Egu O. – Bonnel P.: How comparable are origin-destination matrices estimated from automatic fare collection, origin-destination surveys and household travel survey? An empirical investigation in

- Lyon, Transportation Research Part A: Policy and Practice 2020/138 267-282  
<https://doi.org/10.1016/j.tra.2020.05.021>
- [21] Doblas J. – Benítez F. G.: An approach to estimating and updating origin–destination matrices based upon traffic counts preserving the prior structure of a survey matrix, Transportation Research Part B: Methodological 2005/39 565-591 <https://doi.org/10.1016/j.trb.2004.06.006>
- [22] Lo H-P. – Chan C-P.: Simultaneous estimation of an origin–destination matrix and link choice proportions using traffic counts, Transportation Research Part A: Policy and Practice 2003/37 771-788 [https://doi.org/10.1016/S0965-8564\(03\)00048-X](https://doi.org/10.1016/S0965-8564(03)00048-X)
- [23] Horváth B. – Horváth R.: Real network test of an iterative origin-destination matrix estimator in urban public transport, 2014 18th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia 2014. október 17-19. <https://doi.org/10.1109/ICSTCC.2014.6982502>
- [24] Horváth B.: Uncertainty of the od matrix's estimation in urban public transport, 6th Transport Research Arena, Varsó 2016. április 18-21. <https://doi.org/10.1016/j.trpro.2016.05.137>
- [25] Oszter V.: Long-distance travel patterns of rail users in Hungary - Findings of the new countrywide Origin-Destination (O-D) Survey, Transportation Research Procedia 2020/48 281-2843  
<https://doi.org/10.1016/j.trpro.2020.08.235>
- [26] Pintér J. - Rappai G.: A mintavételi tervek készítésének néhány gyakorlati megfontolása, Marketing & Menedzsment, 2001/35/4 4-10